

The architecture of human memory: insights from human single-neuron recordings

Ueli Rutishauser^{1,2,3}, Leila Reddy^{4,5}, Florian Mormann⁶, Johannes Sarnthein^{7,8}

¹ Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, U.S.A.

² Center for Neural Science and Medicine, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles CA, U.S.A.

³ Division of Biology and Biological Engineering, California Institute of Technology, Pasadena CA, U.S.A.

⁴ Université de Toulouse, UPS, Centre de Recherche Cerveau et Cognition, 31052 Toulouse, France

⁵ CNRS, CerCo, 31052 Toulouse, France

⁶ Department of Epileptology, University of Bonn, Germany

⁷ Klinik für Neurochirurgie, Universitäts Spital und Universität Zürich, Switzerland

⁸ Neuroscience Center Zurich, Universität Zürich und ETH Zürich, Switzerland

Corresponding author: Ueli Rutishauser, ueli.rutishauser@cshs.org

Abstract

Deciphering the mechanisms of human memory is a central goal of neuroscience, both from the point of view of the fundamental biology of memory and for its translational relevance. Here, we review some contributions that recordings from neurons in humans implanted with electrodes for clinical purposes have made towards this goal. Recordings from the medial temporal lobe, including the hippocampus, reveal the existence of two classes of cells: those encoding highly selective and invariant representations of abstract concepts, and memory-selective cells whose activity is related to familiarity and episodic retrieval. Insights derived from observing these cells in behaving humans include that semantic representations are activated before episodic representations, that memory content and memory strength are segregated, and that the activity of both types of cells is related to subjective awareness as expected from a substrate for declarative memory. Visually selective cells can remain persistently active for several seconds, thereby revealing a cellular substrate for working memory in humans. An overarching insight is

that the neural code of human memory is interpretable at the single-neuron level. Jointly, intracranial recording studies are starting to reveal aspects of the building blocks of human memory at the single-cell level. This work establishes a bridge to cellular-level work in animals on the one hand, and the extensive literature on non-invasive imaging in humans on the other hand. More broadly, this work is a step towards a detailed mechanistic understanding of human memory that is needed to develop therapies for human memory disorders.

Introduction

How is it that reminiscing about a salient autobiographical event results in the phenomenological experience of ‘remembering’ (Tulving, 2002) that is accompanied by a multitude of associated memories of what we were feeling at the time, what we were smelling, and what someone said to us? Unique insight into these fundamental questions has been gained by recording from single neurons in human patients implanted with electrodes for clinical purposes (Engel et al., 2005; Suthana and Fried, 2012; Johnson and Knight, 2015; Rutishauser, 2019). This technique allows investigators to observe the activity of individual neurons in areas of the human brain critical to memory (Fried et al., 2014) while patients perform tasks in which they report their subjective experience such as with what confidence they believe a face shown is familiar. This work is starting to reveal the building blocks for human memory (Quiroga, 2012; Quiroga, 2019; Rutishauser, 2019) and thereby identifies features of memory organization that may be uniquely human and others that are shared with other species, thereby building a bridge to the broader literature.

In this review, we highlight advances in our understanding of human memory derived from invasive single-neuron recordings in the human hippocampus, amygdala, parahippocampal gyrus, and entorhinal cortex. These parts of the medial temporal lobe (MTL) (Squire et al., 2004; Eichenbaum et al., 2007) are critical for forming and retrieving episodic memories of the ‘where, what, when’ of past experience. The MTL is also critical for forming and utilizing semantic memories of facts about the world (Manns et al., 2003), which constitute the features based on which episodic memories are constructed. Using this approach, different types of neurons have been identified that, for example, signal aspects of memory content and strength. The trial-by-trial variability in the responses of these neurons explains behavior, such as

whether a memory was retrieved with high or low confidence (Rutishauser et al., 2015) or whether a stimulus was consciously perceived or not (Reber et al., 2017). An overarching insight derived from the work reviewed is that features of both semantic and episodic memories are encoded at the level of individual neurons in humans. Indeed, listening to a single such neuron can be highly informative about the content of memory and/or visual awareness (Rey et al., 2014; Reber et al., 2017; Staresina et al., 2019). This finding is contrary to what would be expected from a fully distributed code (Rogers and McClelland, 2014; Yuste, 2015). Instead, this work reveals that, in humans, the canonical coding principle of the single unit (Barlow, 1972) that has been used to understand sensory representation is applicable to memory.

In describing these studies, we also highlight the utility of behavioral reports. Declarative memories are memories whose content rises into awareness so that their existence can be ‘declared’, a feature of memory that can be studied directly in humans. The importance of the phenomenological experience of retrieving an episodic memory (i.e. ‘mental time travel’) (Boyer, 2008; Suddendorf et al., 2009) for human cognition is amply demonstrated by the profound effect of its loss due to lesions or neurodegenerative disorders. Recording from single neurons in humans offers a unique opportunity to decipher the underlying mechanisms so that, eventually, we might be able to develop new treatment modalities for these dreaded disorders of the human mind.

Building blocks of episodic memories – memory-selective cells

Neurons within the human MTL signal whether stimuli are novel or familiar (Fried et al., 1997; Rutishauser et al., 2006; Viskontas et al., 2006; Rutishauser et al., 2008), thereby revealing

a single-neuron substrate for this core feature of episodic memory. Characterizing these neurons in detail has been a major focus for several reasons. First, a given event happens only once, thus requiring rapid plasticity. Second, successful detection of novelty is a pre-requisite for encoding, thereby allowing novelty signals (Knight, 1996) to serve as signals that indicate encoding. Third, familiarity is graded: a stimulus can either be vaguely familiar with no associated attributes or be highly familiar with many associated attributes (recollection). Fourth, for a memory to be declarative, subjects need to be able to declare the presence or absence of a memory. Thus, a neuronal novelty/familiarity-related response would need to predict awareness or subjective confidence to be relevant for declarative memory. Single-neuron correlates of many of these properties have now been identified.

The response of memory-selective cells differentiates between novel and familiar stimuli. Such cells have been identified across the MTL, including the hippocampus, amygdala, entorhinal cortex, and parahippocampal cortex. Two types of memory-selective cells have been identified: a novelty type and a familiarity type, which increase their response relative to baseline for novel and familiar stimuli, respectively (Fig. 1A). The responses of these cells are compatible with the four criteria outlined above. First, they express single trial learning (Rutishauser et al., 2006; Kaminski et al., 2018). Second, the phase-locking to theta oscillations of the novelty type of memory-selective cells is indicative of encoding success or failure (Rutishauser et al., 2010). Third, their response is graded, with responses accompanied by both familiarity and successful retrieval of an associated stimulus larger than those that are accompanied only by familiarity but no retrieval of a previously learned association (Fig. 1D). This effect has been demonstrated both for cued recall of paired items (Staresina et al., 2019; Derner et al., 2020) and for cued spatial source recall (Rutishauser et al., 2008). Fourth, the response of memory-selective cells

differentiates between otherwise identical correct decisions that are made either with high or low subjective confidence (Fig. 1B) (Rutishauser et al., 2015). Together, this shows that the response of memory-selective cells in the human MTL are likely a reflection of episodic memory. A key next goal is to determine how the signal carried by memory-selective cells is computed and where the plasticity expressed by these cells occurs.

The response properties of memory-selective cells provide constraints on how the signal they carry might be computed. First, their response latency is long: their response differs between novel and familiar stimuli starting at ~450 ms after stimulus onset. Relative to the response latencies of neurons in sensory neocortex, this is slow and too late for behaviors such as basic categorization (Thorpe et al., 1996). Instead, this long latency suggests that the signal carried by memory-selective cells is related to declarative memory, which operates at a significantly slower timescale. Supporting this view, evoked potentials recorded with scalp EEG that are related to declarative memory (such as the N400 or the late negative component) occur with similarly long latencies (Mormann et al., 2005). Second, most neurons with familiarity/novelty responses in the MTL are not tuned to memory content (Fig. 1C).

In addition to the encoding of novelty/familiarity in recognition memory, memory-selective responses have also been observed during cued retrieval in associative memory tasks (Fig. 1D). In this task, a type of memory-selective cell was identified that signaled whether a stimulus associated with a cue was retrieved or not, but not what the identity of the retrieved item was (Staresina et al., 2019). An overall conclusion from these recordings is that the human MTL contains an orthogonal code for memory content and familiarity (Rutishauser et al., 2015; Staresina et al., 2019; Derner et al., 2020).

Building blocks of episodic memories – semantically invariant concept cells

The response of many neurons in the human amygdala, hippocampus, entorhinal cortex, and parahippocampal cortex is related to the high-level visual interpretation of stimuli. A continuum of such cells has been discovered (Fig. 2A-C shows examples), from relatively broadly tuned, category cells to more sparsely tuned cells signaling semantic concepts, all the way to highly sparse and invariant concept cells that respond whenever a picture shows a particular individual. We refer to such cells jointly as “visually selective”, but note that this is not meant to imply that these cells do not also respond to other modalities (Quiroga et al., 2009). The most broadly tuned type of visually-selective cells is category neurons (Fig. 2C), which respond every time an image belonging to a particular broad visual category is shown (Kreiman et al., 2000). Examples of category neurons are neurons responding to faces (Minxha et al., 2017), animals (Mormann et al., 2011), outdoor scenes (Mormann et al., 2017), fruits (Minxha et al., 2020), and various other object categories (Reber et al., 2019a). The response of category cells is insensitive to whether a stimulus is novel or familiar as well as to whether a stimulus was retrieved with high or low confidence (Rutishauser et al., 2015) (note that this does not exclude short-lasting repetition suppression (Pedreira et al., 2010; Minxha et al., 2020)). At the other extreme of selectivity and invariance are ‘concept cells’, which become active briefly only when a particular high-level concept is shown (Quiroga et al., 2005). Such cells can develop relatively rapidly as demonstrated by cells that respond to experimenters previously unknown to the subject (Fig. 2B). Most concept cells respond to concepts personally known to the person the neuron was recorded from (Viskontas et al., 2009). While the response of concept cells is sparse, these cells are not rare: it has been estimated that 2–5 million MTL cells respond to a given concept (Waydo et al., 2006).

The response of visually-selective cells is indicative of whether a given stimulus has entered conscious awareness of the subject as declared by subjective report (Kreiman et al., 2002; Quiroga et al., 2008; Quian Quiroga et al., 2014; Rey et al., 2014; Reber et al., 2017). The more anterior in the MTL a concept cell is located, the more likely its activity is to co-vary with awareness (Reber et al., 2017; Fu and Rutishauser, 2018). The view that responses of visually-selective cells are related to the current content of awareness is in addition supported by the finding that in dense scenes, category cells signal the categorical identity of the currently attended stimulus, both when attention is measured overtly by eye movements or covertly redeployed (Minxha et al., 2017). Visually-selective cells thus appear well-suited to encode declarative memories, which are memories which have risen into awareness.

The time that elapses between onset of a stimulus and a differential response of visually-selective cells is approximately 300ms (Mormann et al., 2008; Rutishauser et al., 2015). The latency till a an individual cell starts responding is inversely proportional to its selectivity, with more selective responses occurring later (Mormann et al., 2008). This is late compared to responses in sensory areas measured using similar techniques (Self et al., 2016), but earlier than that of MS cells (discussed above). Also, this long latency is what is expected from a stimulus that has risen into awareness (Mashour et al., 2020) and is concordant with the known response latencies of memory-related ERPs (Mormann et al., 2005).

These data have given rise to the hypothesis that visually-selective cells represent semantic memories (Rutishauser, 2019), which in turn are the building blocks of episodic memories (Quiroga, 2012). Further evidence for this proposal is provided by population-level analysis of a large set of single neurons (Reber et al., 2019a). This analysis revealed that the response of visually-selective cells generalized between instances of a given semantic concept.

Whereas some semantic concepts represented by cells were relatively broad ('food items'), others were more narrow ('cheese', see Fig. 2A). Indeed, whether a given cell responds to a given stimulus or not could be predicted with high accuracy from the semantic label of a stimulus (Fig. 2A). Notably, the response of cells was not 'all-or-nothing'. Rather, the response was best described as a tuning curve in semantic space with graded responses to different instances (stimuli) from a given semantic category (for example, a cell that responds with a different magnitude to a picture of shorts and a red jacket, see Fig. 2A top row). This is a critical new insight that supports the hypothesis that visually-selective cells represent semantic memories, and as such are a core part of the engram. This view, however, is challenged by two recent studies describing responses of concept neurons to different person identities as binary rather than graded (Rey et al., 2018; Rey et al., 2020). Future work is needed to decipher a better understanding of the regimes under which visually-selective cell responses can be conceptualized as binary vs. graded, in particular whether these constitute functionally different subgroups of cells and whether these two types of cells have a differential role in declarative memory.

Encoding of associations

Memory associates different aspects of experience with each other. For example, an episodic memory might be that I met person X at place Y at time Z. Visually-selective cells by themselves represent an encoding of associations: if two stimuli are conceptually related, they activate the same cell (i.e., all stimuli that belong to the same semantic concept). This property also holds for more sparsely tuned cells: when such a cell responds to several seemingly unrelated stimuli (such as two different basketball players that play on the same team), these

stimuli are indeed conceptually related as revealed by an analysis of how often the two semantic concepts co-occur (De Falco et al., 2016). This broadening of tuning can be induced experimentally: repeated pairing of a responsive and non-responsive concept leads to a broadening of the tuning of an individual cell, which starts to respond also to the non-response-eliciting concept as a result of learning (Ison et al., 2015). Associative encoding has also been observed in the temporal domain: if a subject learns that a picture of concept X is always followed by that of concept Y, the cell responsive to Y starts firing already in the delay period before image Y appears (Reddy et al., 2015). Of note, the timescale of plasticity in the data discussed above is relatively long (several trials). Thus, these changes occur on the temporal scale expected of semantic but not episodic memories, which rely on single-trial learning. Together, these data show that the representations formed by visually-selective cells are plastic as expected from a semantic representation that supports episodic memory.

A second form of associative memory that has been studied is cued retrieval: after learning pairs (such as face-name pairs), one of the items (i.e the face) is shown and the subjects attempts to retrieve the paired stimulus (i.e. the name). Recordings in the entorhinal cortex (Staresina et al., 2019) reveal cells which signal the content of the retrieved stimulus (the name in above example) (see Fig. 2D-E). Reinstatement of the memory content retrieved after the cue is shown is sufficiently strong to allow decoding of the recollected memory. Indeed, earlier work has revealed that reinstatement extends beyond neurons specifically tuned to retrieved memory content: the neural state ('neural context') that was present at the time of encoding tends to be reinstated during successful recollection (Jang et al., 2017). Together, this body of work reveals a single-neuron correlate of reinstatement, thereby providing an empirical basis for a key theory about how episodic memory is organized (Howard et al., 2005; Polyn et al., 2009).

Persistent activity as a mechanism for working memory

Some MTL cells can remain persistently active for several seconds while subjects actively hold the cell's preferred stimulus in working memory (WM), a finding first reported in 2017 independently by two groups (Kaminski et al., 2017; Kornblith et al., 2017). Decoding from such persistently active cells allows the decoding of the content of working memory (Kaminski et al., 2020). Also, the cells' activity decreases as a function of memory load, as expected from a representation of short-term memory (Kaminski et al., 2017), and their response strength predicts whether the memory will later be retrieved successfully (Kornblith et al., 2017). This pattern of activation is different from that seen in a typical experiment, in which these cells show transient activation lasting <500 ms after stimulus onset. During WM tasks, persistently active visually-selective cells maintain their selectivity and preference, such that the stimulus to which they are selective during stimulus presentation remains the same for several seconds during maintenance, when no image is shown (Fig. 3A-C). About 50% of concept cells exhibit such continuous content-selective activation. A second type of persistently active cell has been discovered in the hippocampus: cells whose persistent activity increases as a function of load but which is otherwise not content selective (Fig. 3D). Together, these studies reveal that the substrate for long-term memories can also support WM. More broadly, this work reveals a cellular substrate for human WM: that of persistently active cells. The electrophysiological properties of these human cells appear much like those reported since the earliest such studies in the macaque frontal lobe (Fuster and Alexander, 1971) or recently in rodent MTL (Masuda et al., 2020), except that here their tuning is for high-level semantic concepts rather than spatial locations.

It might appear paradoxical that cells in the human MTL would support WM since the MTL is not strictly necessary for WM. While this is true for many tasks (Squire et al., 2004), under more challenging circumstances such as interference from distractors or high memory load, humans with MTL lesions do exhibit WM deficits (Jeneson and Squire, 2012). Furthermore, there is evidence that the MTL may support maintenance of visuo-spatial relations between objects at sufficient levels of precision in working memory (Alvarez and Cavanagh, 2004). Evidence for a role of the MTL in WM maintenance can also be seen at the field potential level: both sustained theta-band activity in the MTL and theta-mediated PFC-MTL interactions are prominent in the MTL during WM maintenance (Johnson et al., 2018). Based on this, we hypothesize that the persistently active MTL cells in humans support WM under these more challenging circumstances in which purely sensory representations are no longer sufficient to maintain the memory. The association of MTL visually-selective cell activity with conscious awareness allows a further hypothesis: that these cells represent WM content that is presently active in awareness, an active form of WM that has long been recognized to be distinct from other forms of WM such as activated long-term memory in theoretical models of WM (Cowan, 1988; Kaminski and Rutishauser, 2020). While many aspects of this relationship remain to be investigated, this identification of a cellular substrate of WM that is semantically selective offers many opportunities to directly test and refine these important cognitive models.

From memories to decisions

How are representations of memories in the MTL used for making decisions? Neuroimaging reveals several cortical areas that might carry memory-related choice signals (Simons and Spiers, 2003; Rugg and Vilberg, 2013), including in particular the medial frontal

cortex (MFC) and left posterior parietal cortex (Wagner et al., 2005) [our focus here is on the MFC, see (Rutishauser et al., 2018) for results on the PPC]. In rodents, the MFC receives monosynaptic connections from hippocampal neurons, and these connections are causally relevant for making memory-based decisions during spatial navigation (Spellman et al., 2015; Tamura et al., 2017).

Neurons within the human pre-supplementary motor area (preSMA) and dorsal anterior cingulate cortex (dACC) signal decisions about whether a stimulus is novel or familiar in a recognition memory task (Minxha et al., 2020), a role distinct from the traditionally studied role of the MFC in cognitive control and performance monitoring (Fu et al., 2019). The neurons signaled these decisions in an abstract format irrespective of motor output modality, indicating that their response is a reflection of an abstract choice representation rather than motor actions. The neurons that signaled memory-based decisions were largely distinct from those that signaled perceptual decisions. Also, neurons carrying memory-based choice signals flexibly adjusted their phase-locking with hippocampal theta oscillations: they increased their phase-locking only when a decision required access to memory. Finally, memory-based choice cells signal the decision about the memory rather than the memory itself: in cases of errors, the activity of the cells predicted the subject's choice irrespective of whether it was true or false. Together, this work reveals a mechanism by which cortical decision processes flexibly interact with the MTL.

Outlook: new technology to record human single neurons?

The recordings discussed above rely on novel types of hybrid electrodes (Cash and Hochberg, 2015) that allow recording from microwires in patients with epilepsy undergoing

placement of depth electrodes for localization of their seizure focus (Babb et al., 1973). The majority of studies use variants of the Behnke-Fried electrode (Fried et al., 1999), which has been in routine use for many years. This electrode consists of 8-9 microwires protruding from the tip of a depth electrode. These wires exit the carrier in an uncontrolled fashion, thereby preventing their use as tetrodes to improve spike sorting or to calculate current source density. A notable development in this direction is tetrodes in humans (Despouy et al., 2020), which promise improved yield and isolation quality. A key next milestone will be to adapt and safely deploy newer high-density recording methods that will make it possible to simultaneously record from hundreds or thousands of neurons (Juavinett et al., 2019). Taking this major step is a key challenge for the field because of the considerable effort needed to demonstrate the safety of this technology for human use.

Conclusion

In this review, we discussed two broad classes of cells in the human MTL: those that selectively respond to specific subsets of stimuli and those that differentiate between different aspects of memory. The latter group is not content-selective: they signal, for example, whether a stimulus is novel or familiar, or whether an associated attribute was retrieved or not. The activity of these cells changes after a single exposure and predicts aspects of declared subjective human experience, such as confidence and the experience of recollection. In models of declarative memory (Yonelinas, 2001; Wixted, 2007), decision variables that signal ‘memory strength’ that are otherwise content-free are a common assumption. The signals carried by memory-selective cells in the MTL share many of the properties predicted by such models, thereby enabling a close interaction between theory and circuit-level experiments.

Are visually-selective cells in the human MTL a kind of a ‘place cell’ (as found in rodents) in the cognitive space formed by our semantic memories (Lisman et al., 2017) ? While there are many striking similarities (see (Quiroga, 2012) for details), there are also salient differences that call this analogy into question. A remarkable feature of visually-selective cells in humans is their context-independence, including being active during the very first exposure to a novel stimulus. This is unlike place cells in rodents and non-human primates, which are highly dependent on context and behavior (Anderson and Jeffery, 2003; Courellis et al., 2019). For example, place cell firing depends on direction of movement, and modifications of the environment or task goals can lead to complete remapping. While this question needs more systematic study, there are critical differences between the properties of these two types of cells.

Whole-brain fMRI studies have revealed remarkable patterns of semantic tuning that tile the human cortex (Huth et al., 2016). These patterns are seen at different levels of semantic abstraction, which is reminiscent of how visually-selective cells respond to semantic concepts. An intriguing open question is whether these two findings, made at very different scales of observation, are an expression of the same phenomena. Future work is needed to determine whether the trials on which a semantically tuned cell in the MTL responds are the same trials as those on which the corresponding parts of cortex become active as assessed by fMRI. If so, this might indicate that VS cells in the MTL that are semantically tuned are cells that ‘index’ or ‘point’ to cortical patterns of activity that represent memory content, a mode of hippocampal operation long theorized (Teyler and Rudy, 2007).

Lastly, it is imperative to point out that without work performed in animal models, the rare and select opportunities to record invasively from humans (which gave rise to all the results discussed here) could not be used meaningfully. Indeed, these two approaches are

complementary: a key role for human invasive experiments is to establish a bridge between species by establishing the relevance of certain phenomena for human memory, behavior, and disease. As a result of such links, the identified mechanisms can then be investigated in detail at the molecular-and cellular level in model animals (Kandel et al., 2014) in ways that is impossible or unethical in humans but which at the same time is of direct relevance because of the established links between the species.

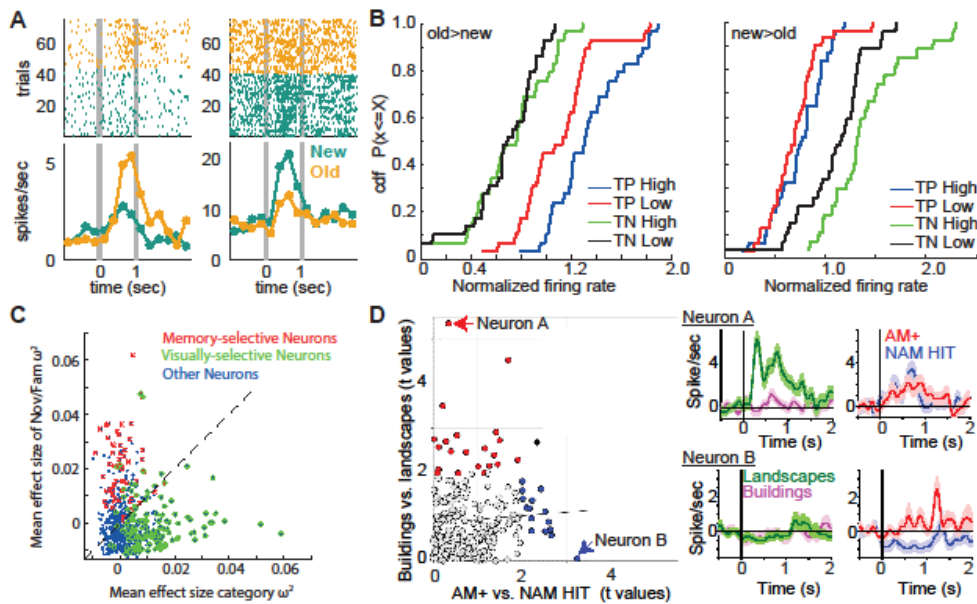


Figure 1: Encoding episodic memory strength and retrieval mode. (A) Example memory-selective cells. Left: cell that fires most to familiar (old) stimuli. Right: cell that fires most to novel (new) stimuli. $t=0$ is stimulus onset. Top shows raster plots, with each row an individual trial and each dot an action potential. Bottom shows average firing rate across all trials binned 250ms bins. (B) Activity of both types of memory-selective cells scale with confidence for their preferred but not their non-preferred stimulus (either old [left] or new [right] stimuli). TP=true positive (old stimuli recognized correctly) shown as a cumulative density function (cdf). TN=true negative (new stimuli recognized correctly). (C) Orthogonal tuning to either visual category (x axis) or familiarity (y axis) during recognition memory. Each datapoint is a neuron. Red and green colors identify neurons that are primarily influenced by novelty/familiarity or category, respectively. (D) Orthogonal tuning to either visual category (y axis) or successful retrieval vs. familiarity only (x axis). Shown are two example neurons (marked), one only visually selective (Neuron A), the other only differentiating successful retrieval from familiarity alone (Neuron B). Each datapoint is a neuron. Red and blue identify neurons that are primarily influenced by visual identity and success of retrieval of the associated stimulus, respectively. (A) Adapted from (Faraut et al., 2018). (B-C) Adapted from (Rutishauser et al., 2015). (D) Adapted from (Staresina et al., 2019).

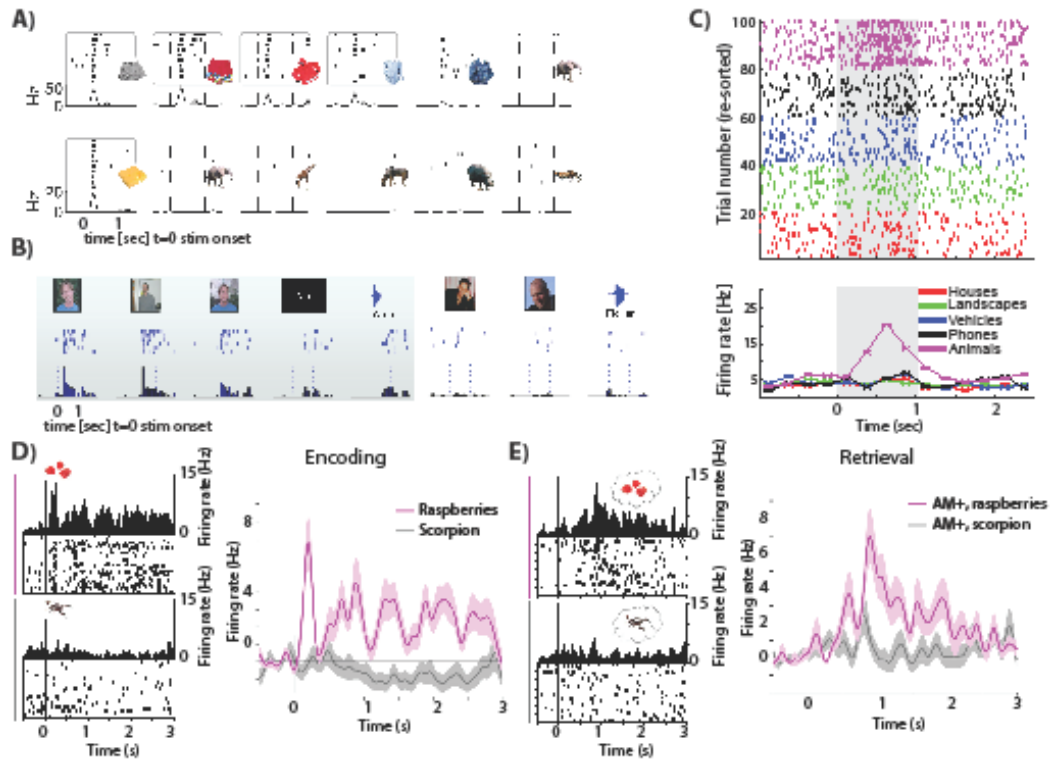


Figure 2: Abstract semantic representations and their role in episodic retrieval. (A-C) Example of visually selective neurons with tuning compatible with semantic representations. (A) Top: visually selective neuron responding to many different images showing clothes (from hippocampus). Bottom: Visually selective neuron only responding to a single image of a food item (from amygdala). (B) Invariant multimodal concept neuron that responds only to images and written and spoken name of an experimenter. (C) Visually selective category neuron. $T=0$ is stimulus onset. Top shows raster plot showing the spiking response to different stimulus categories (indicated by color), bottom shows average firing rate across trials in bins of 250ms. (D) Visually selective neuron in the entorhinal cortex with a response to ‘raspberries’ but not to ‘scorpion’ during encoding. (E) Same neuron as shown in (D), but now during cued retrieval. This neuron increased its firing selectively only if associated to-be retrieved image was ‘raspberries’. This is a signature of content-specific reinstatement during cued retrieval. (A) adapted from (Reber et al., 2019b). (B) adapted from (Quian Quiroga et al., 2009). (C) adapted from (Rutishauser, 2019). (D-E) adapted from (Staresina et al., 2019).

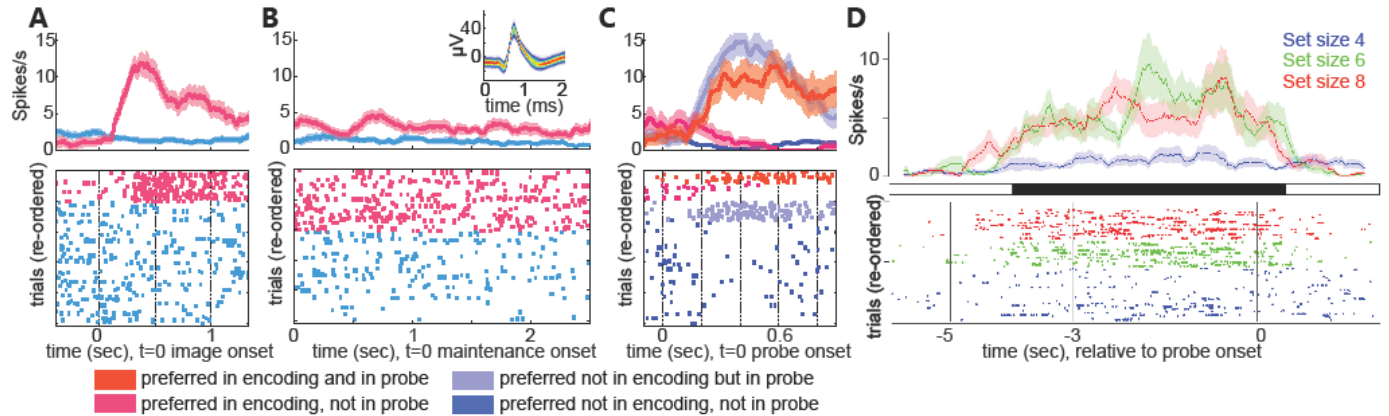


Figure 3: Persistently active cells as a substrate for working memory. (A-C) Example of a visually selective neuron recorded from the amygdala during a working memory task with pictures as items. Shown is the encoding (A), maintenance (B), and probe (C) period. The firing rate of this neuron increases starting during encoding when its preferred stimulus is shown (magenta) and maintains this activity during the maintenance period if its preferred image is held in memory (B). (B) inset shows the mean waveform of this neuron. (D) Example neuron in the hippocampus recorded during a working memory task with letters as items. This neuron increases its activity during the maintenance period as a function of load but not memory content. In all panels, bottom shows raster plots showing the spiking response in individual trials (one per row), with identity of the trial marked by color. Top shows average firing rate as a function of time across all trials as a function of time. (A-C) adapted from (Kaminski et al., 2020). (D) Adapted from (Boran et al., 2019).

References

- Alvarez GA, Cavanagh P (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol Sci* 15:106-111.
- Anderson MI, Jeffery KJ (2003) Heterogeneous modulation of place cell firing by changes in context. *J Neurosci* 23:8827-8835.
- Babb TL, Carr E, Crandall PH (1973) Analysis of extracellular firing patterns of deep temporal lobe structures in man. *Electroencephalogr Clin Neurophysiol* 34:247-257.
- Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1:371-394.
- Boran E, Fedele T, Klaver P, Hilfiker P, Stieglitz L, Grunwald T, Sarnthein J (2019) Persistent hippocampal neural firing and hippocampal-cortical coupling predict verbal working memory load. *Sci Adv* 5:eaav3687.
- Boyer P (2008) Evolutionary economics of mental time travel? *Trends Cogn Sci* 12:219-224.
- Cash SS, Hochberg LR (2015) The emergence of single neurons in clinical neurology. *Neuron* 86:79-91.
- Courellis HS, Nummela SU, Metke M, Diehl GW, Bussell R, Cauwenberghs G, Miller CT (2019) Spatial encoding in primate hippocampus during free navigation. *PLoS Biol* 17:e3000546.
- Cowan N (1988) Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol Bull* 104:163-191.
- De Falco E, Ison MJ, Fried I, Quiñero R (2016) Long-term coding of personal and universal associations underlying the memory web in the human brain. *Nat Commun* 7:13408.
- Derner M, Dehnen G, Chaieb L, Reber TP, Borger V, Surges R, Staresina BP, Mormann F, Fell J (2020) Patterns of single-neuron activity during associative recognition memory in the human medial temporal lobe. *Neuroimage* 221:117214.
- Despouy E, Curot J, Reddy L, Nowak LG, Deudon M, Sol JC, Lotterie JA, Denuelle M, Maziz A, Bergaud C, Thorpe SJ, Valton L, Barbeau EJ (2020) Recording local field potential and neuronal activity with tetrodes in epileptic patients. *J Neurosci Methods* 341:108759.
- Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123-152.
- Engel AK, Moll CK, Fried I, Ojemann GA (2005) Invasive recordings from the human brain: clinical insights and beyond. *Nat Rev Neurosci* 6:35-47.
- Faraut MCM, Carlson AA, Sullivan S, Tudusciuc O, Ross I, Reed CM, Chung JM, Mamelak AN, Rutishauser U (2018) Dataset of human medial temporal lobe single neuron activity during declarative memory encoding and recognition. *Scientific Data* 5:180010.
- Fried I, MacDonald KA, Wilson CL (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18:753-765.
- Fried I, Rutishauser U, Cerf M, Kreiman G (2014) *Single Neuron Studies of the Human Brain: Probing Cognition*. Boston: MIT Press.
- Fried I, Wilson CL, Maidment NT, Engel J, Jr., Behnke E, Fields TA, MacDonald KA, Morrow JW, Ackerson L (1999) Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients. Technical note. *J Neurosurg* 91:697-705.
- Fu Z, Rutishauser U (2018) Single-Neuron Correlates of Awareness during Attentional Blinks. *Trends Cogn Sci* 22:5-7.
- Fu Z, Wu DJ, Ross I, Chung JM, Mamelak AN, Adolphs R, Rutishauser U (2019) Single-Neuron Correlates of Error Monitoring and Post-Error Adjustments in Human Medial Frontal Cortex. *Neuron* 101:165-177 e165.
- Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652-654.

Howard MW, Fotedar MS, Datey AV, Hasselmo ME (2005) The temporal context model in spatial navigation and relational learning: toward a common explanation of medial temporal lobe function across domains. *Psychol Rev* 112:75-116.

Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532:453-458.

Ison MJ, Quiroga R, Fried I (2015) Rapid Encoding of New Memories by Individual Neurons in the Human Brain. *Neuron* 87:220-230.

Jang AI, Wittig JH, Jr., Inati SK, Zaghoul KA (2017) Human Cortical Neurons in the Anterior Temporal Lobe Reinstatement Spiking Activity during Verbal Memory Retrieval. *Curr Biol* 27:1700-1705 e1705.

Jones A, Squire LR (2012) Working memory, long-term memory, and medial temporal lobe function. *Learn Mem* 19:15-25.

Johnson EL, Knight RT (2015) Intracranial recordings and human memory. *Curr Opin Neurobiol* 31:18-25.

Johnson EL, Adams JN, Solbakk AK, Endestad T, Larsson PG, Ivanovic J, Meling TR, Lin JJ, Knight RT (2018) Dynamic frontotemporal systems process space and time in working memory. *PLoS Biol* 16:e2004274.

Juavinett AL, Bekheet G, Churchland AK (2019) Chronically implanted Neuropixels probes enable high-yield recordings in freely moving mice. *Elife* 8.

Kaminski J, Rutishauser U (2020) Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Ann N Y Acad Sci* 1464:64-75.

Kaminski J, Brzezicka A, Mamelak AN, Rutishauser U (2020) Combined Phase-Rate Coding by Persistently Active Neurons as a Mechanism for Maintaining Multiple Items in Working Memory in Humans. *Neuron* 106:256-264 e253.

Kaminski J, Sullivan S, Chung JM, Ross IB, Mamelak AN, Rutishauser U (2017) Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat Neurosci* 20:590-601.

Kaminski J, Mamelak AN, Birch K, Mosher CP, Tagliati M, Rutishauser U (2018) Novelty-Sensitive Dopaminergic Neurons in the Human Substantia Nigra Predict Success of Declarative Memory Formation. *Curr Biol* 28:1333-1343 e1334.

Kandel ER, Dudai Y, Mayford MR (2014) The molecular and systems biology of memory. *Cell* 157:163-186.

Knight R (1996) Contribution of human hippocampal region to novelty detection. *Nature* 383:256-259.

Kornblith S, Quiroga R, Koch C, Fried I, Mormann F (2017) Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr Biol* 27:1026-1032.

Kreiman G, Koch C, Fried I (2000) Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat Neurosci* 3:946-953.

Kreiman G, Fried I, Koch C (2002) Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc Natl Acad Sci U S A* 99:8378-8383.

Lisman J, Buzsaki G, Eichenbaum H, Nadel L, Ranganath C, Redish AD (2017) Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nat Neurosci* 20:1434-1447.

Manns JR, Hopkins RO, Squire LR (2003) Semantic memory and the human hippocampus. *Neuron* 38:127-133.

Mashour GA, Roelfsema P, Changeux JP, Dehaene S (2020) Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron* 105:776-798.

Masuda A, Sano C, Zhang Q, Goto H, McHugh TJ, Fujisawa S, Itohara S (2020) The hippocampus encodes delay and value information during delay-discounting decision making. *eLife* 9:e52466.

Minxha J, Adolphs R, Fusi S, Mamelak AN, Rutishauser U (2020) Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* 368.

- Minxha J, Mosher C, Morrow JK, Mamelak AN, Adolphs R, Gothard KM, Rutishauser U (2017) Fixations Gate Species-Specific Responses to Free Viewing of Faces in the Human and Macaque Amygdala. *Cell Rep* 18:878-891.
- Mormann F, Fell J, Axmacher N, Weber B, Lehnertz K, Elger CE, Fernandez G (2005) Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus* 15:890-900.
- Mormann F, Kornblith S, Quiroga RQ, Kraskov A, Cerf M, Fried I, Koch C (2008) Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J Neurosci* 28:8865-8872.
- Mormann F, Kornblith S, Cerf M, Ison MJ, Kraskov A, Tran M, Knieling S, Quian Quiroga R, Koch C, Fried I (2017) Scene-selective coding by single neurons in the human parahippocampal cortex. *Proc Natl Acad Sci U S A* 114:1153-1158.
- Mormann F, Dubois J, Kornblith S, Milosavljevic M, Cerf M, Ison M, Tsuchiya N, Kraskov A, Quiroga RQ, Adolphs R, Fried I, Koch C (2011) A category-specific response to animals in the right human amygdala. *Nat Neurosci* 14:1247-1249.
- Pedreira C, Mormann F, Kraskov A, Cerf M, Fried I, Koch C, Quiroga RQ (2010) Responses of human medial temporal lobe neurons are modulated by stimulus repetition. *J Neurophysiol* 103:97-107.
- Polyn SM, Norman KA, Kahana MJ (2009) A context maintenance and retrieval model of organizational processes in free recall. *Psychol Rev* 116:129-156.
- Quian Quiroga R (2019) Plugging in to Human Memory: Advantages, Challenges, and Insights from Human Single-Neuron Recordings. *Cell* 179:1015-1032.
- Quian Quiroga R, Kraskov A, Koch C, Fried I (2009) Explicit encoding of multimodal percepts by single neurons in the human brain. *Curr Biol* 19:1308-1313.
- Quian Quiroga R, Kraskov A, Mormann F, Fried I, Koch C (2014) Single-cell responses to face adaptation in the human medial temporal lobe. *Neuron* 84:363-369.
- Quiroga RQ (2012) Concept cells: the building blocks of declarative memory functions. *Nat Rev Neurosci* 13:587-597.
- Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435:1102-1107.
- Quiroga RQ, Mukamel R, Isham EA, Malach R, Fried I (2008) Human single-neuron responses at the threshold of conscious recognition. *Proc Natl Acad Sci U S A* 105:3599-3604.
- Reber TP, Faber J, Niediek J, Bostrom J, Elger CE, Mormann F (2017) Single-Neuron Correlates of Conscious Perception in the Human Medial Temporal Lobe. *Curr Biol* 27:2991-2998 e2992.
- Reber TP, Bausch M, Mackay S, Bostrom J, Elger CE, Mormann F (2019a) Representation of abstract semantic knowledge in populations of human single neurons in the medial temporal lobe. *PLoS Biol* 17:e3000290.
- Reber TP, Bausch M, Mackay S, Bostrom J, Elger CE, Mormann F (2019b) Representation of Abstract Semantic Knowledge in Populations of Human Single Neurons in the Medial Temporal Lobe. *PLoS Biol* in press.
- Reddy L, Poncet M, Self MW, Peters JC, Douw L, van Dellen E, Claus S, Reijneveld JC, Baayen JC, Roelfsema PR (2015) Learning of anticipatory responses in single neurons of the human medial temporal lobe. *Nat Commun* 6:8556.
- Rey HG, Fried I, Quian Quiroga R (2014) Timing of single-neuron and local field potential responses in the human medial temporal lobe. *Curr Biol* 24:299-304.
- Rey HG, De Falco E, Ison MJ, Valentin A, Alarcon G, Selway R, Richardson MP, Quian Quiroga R (2018) Encoding of long-term associations through neural unitization in the human medial temporal lobe. *Nat Commun* 9:4372.

- Rey HG, Gori B, Chaure FJ, Collavini S, Blenkmann AO, Seoane P, Seoane E, Kochen S, Quian Quiroga R (2020) Single Neuron Coding of Identity in the Human Hippocampal Formation. *Curr Biol* 30:1152-1159 e1153.
- Rogers TT, McClelland JL (2014) Parallel Distributed Processing at 25: further explorations in the microstructure of cognition. *Cogn Sci* 38:1024-1077.
- Rugg MD, Vilberg KL (2013) Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol* 23:255-260.
- Rutishauser U (2019) Testing Models of Human Declarative Memory at the Single-Neuron Level. *Trends Cogn Sci* 23:510-524.
- Rutishauser U, Mamelak AN, Schuman EM (2006) Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron* 49:805-813.
- Rutishauser U, Schuman EM, Mamelak AN (2008) Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proc Natl Acad Sci U S A* 105:329-334.
- Rutishauser U, Ross IB, Mamelak AN, Schuman EM (2010) Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* 464:903-907.
- Rutishauser U, Aflalo T, Rosario ER, Pouratian N, Andersen RA (2018) Single-Neuron Representation of Memory Strength and Recognition Confidence in Left Human Posterior Parietal Cortex. *Neuron* 97:209-220 e203.
- Rutishauser U, Ye S, Koroma M, Tudusciuc O, Ross IB, Chung JM, Mamelak AN (2015) Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nat Neurosci* 18:1041-1050.
- Self MW, Peters JC, Possel JK, Reithler J, Goebel R, Ris P, Jeurissen D, Reddy L, Claus S, Baayen JC, Roelfsema PR (2016) The Effects of Context and Attention on Spiking Activity in Human Early Visual Cortex. *PLoS Biol* 14:e1002420.
- Simons JS, Spiers HJ (2003) Prefrontal and medial temporal lobe interactions in long-term memory. *Nat Rev Neurosci* 4:637-648.
- Spellman T, Rigotti M, Ahmari SE, Fusi S, Gogos JA, Gordon JA (2015) Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature* 522:309-314.
- Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279-306.
- Staresina BP, Reber TP, Niediek J, Bostrom J, Elger CE, Mormann F (2019) Recollection in the human hippocampal-entorhinal cell circuitry. *Nat Commun* 10:1503.
- Suddendorf T, Addis DR, Corballis MC (2009) Mental time travel and the shaping of the human mind. *Philos Trans R Soc Lond B Biol Sci* 364:1317-1324.
- Suthana N, Fried I (2012) Percepts to recollections: insights from single neuron recordings in the human brain. *Trends Cogn Sci* 16:427-436.
- Tamura M, Spellman TJ, Rosen AM, Gogos JA, Gordon JA (2017) Hippocampal-prefrontal theta-gamma coupling during performance of a spatial working memory task. *Nat Commun* 8:2182.
- Teyler TJ, Rudy JW (2007) The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus* 17:1158-1169.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520-522.
- Tulving E (2002) Episodic memory: From mind to brain. *Annual Review of Psychology* 53:1-25.
- Viskontas IV, Quiroga RQ, Fried I (2009) Human medial temporal lobe neurons respond preferentially to personally relevant images. *Proc Natl Acad Sci U S A* 106:21329-21334.
- Viskontas IV, Knowlton BJ, Steinmetz PN, Fried I (2006) Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *J Cogn Neurosci* 18:1654-1662.
- Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005) Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* 9:445-453.

- Waydo S, Kraskov A, Quian Quiroga R, Fried I, Koch C (2006) Sparse representation in the human medial temporal lobe. *J Neurosci* 26:10232-10234.
- Wixted JT (2007) Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev* 114:152-176.
- Yonelinas AP (2001) Components of episodic memory: the contribution of recollection and familiarity. *Philos Trans R Soc Lond B Biol Sci* 356:1363-1374.
- Yuste R (2015) From the neuron doctrine to neural networks. *Nat Rev Neurosci* 16:487-497.